5e

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/655,325	09/05/2000	Richard Mark Schwartz	99-470	7881

32127 7590 03/22/2004

VERIZON CORPORATE SERVICES GROUP INC.
C/O CHRISTIAN R. ANDERSEN
600 HIDDEN RIDGE DRIVE
MAILCODE HQEO3H14
IRVING, TX 75038

EXAMINER	
PATEL, KINARI M	
ART UNIT	**PAPER NUMBER**
2654	

DATE MAILED: 03/22/2004 5

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/655,325	SCHWARTZ ET AL.	
	Examiner	Art Unit	
	Kinari Patel	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *05 September 2000* .

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-44* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-44* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *05 September 2000* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 1 and 4-14 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nakagawa et al. (WO 00/46788) in view of Zhao (US Patent 5,193,142) and Gales et al. ("State-Based Gaussain Selection in Large Vocabulary Continuous Speech Recognition Using HMM's, IEEE Transactions on Speech and Audio Processing, Vol. 7, No. 2, March 1999).

As per claim 1, Nakagawa et al. disclose a method for recognizing speech, comprising:

receiving an input speech vector (Abstract, Ln. 12-14);

identifying a Gaussian distribution (Abstract, Ln. 1-3);

determining an address from the input speech vector (Abstract, Ln. 25-26);

using the address to retrieve a value for the Gaussian distribution from a table (Abstract, Ln. 15-17 and 20-23);

Nakagawa et al. fail to disclose:

using the address to retrieve a distance value for the Gaussian distribution in a table;

determining a probability of the Gaussian distribution using the distance value; and recognizing the input speech vector based on the determined probability.

The aforementioned features are well known in the art as taught by Zhao and Gales et al. Zhao teaches storing a distance value for Gaussian distributions in a table (Vol. 4, Ln. 34-37 and Ln. 56-58). The distance values can be retrieved in the table instead of the dispersion and average values of Nakagawa et al. Gales et al. teach determining a probability of the Gaussian distribution using the distance value (Page 2, Col. 1, Ln. 46-48 and Col. 2, Ln. 45-50) and recognizing the input speech vector based on the determined probability (Page 1, Col. 1, Ln. 24-26 and Col. 2, Ln. 19-34). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Nakagawa et al. with the steps of Zhao and Gales et al. in order to effectively use Gaussian distributions to model speech.

As per claim 4, Nakagawa et al. as modified by Zhao and Gales et al. teach all the limitations of the method of claim the method of claim 2. Nakawaga et al. fail to disclose the method of claim 2 wherein the determining an address includes: concatenating a code representing the quantized input speech vector and a code representing the Gaussian distribution to form the address for accessing the table. Concatenating a code representing the quantized input speech vector and a code representing the Gaussian distribution to form the address for accessing the table is obvious in the art. Since an input value for the input speech vector is selected from a subset of possible quantized input values and a value is determined for the Gaussian distribution, it is obvious to use those already computed values to form a code to access

a table. Moreover, it is obvious to concatenate any two values to form a code. Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Nakagawa et al. wherein determining an address includes: concatenating a code representing the quantized input speech vector and a code representing the Gaussian distribution to form the address for accessing the table since the obvious way to form a code is to combine values.

As per claim 5, Nakagawa et al. as modified by Zhao and Gales et al. disclose all the limitations of the method of claim 1. Nakagawa further discloses the method of claim 1 wherein the identifying includes identifying a plurality of Gaussian distributions (Abstract, Ln. 20-23).

As per claim 6, Nakagawa et al. as modified by Zhao and Gales et al. teach all the limitations of the method of claim 5. Nakagawa et al. fails to disclose the method of claim 5 further comprising:

repeating the steps of determining an address, using, and determining a probability for each of the Gaussian distributions; and

identifying one or more of the Gaussian distributions with highest probabilities.

The aforementioned features are well known in the art. Nakagawa et al. teach generating a plurality of Gaussian density models (Abstract, Ln. 20-23). Nakagawa et al. further teach determining an address from the input speech vector (Abstract, Ln. 25-26) and describe using the address to retrieve values for the Gaussian distribution from a table (Abstract, Ln. 15-17 and 20-23). Zhao teaches storing a distance value for Gaussian distributions in a table (Vol. 4, Ln. 34-37

and Ln. 56-58). The distance values can be retrieved in the table instead of the dispersion and average values of Nakagawa et al. Gales et al. teach determining a probability of the Gaussian distribution using the distance value (Page 2, Col. 1, Ln. 46-48 and Col. 2, Ln. 45-50) and recognizing the input speech vector based on the determined probability (Page 1, Col. 1, Ln. 24-26 and Col. 2, Ln. 19-34). It is obvious that if determining an address, using, and determining a probability can be performed for one Gaussian distribution, the same steps can be repeated for several or many Gaussian distributions. It is further obvious to identify one or more of the Gaussian distributions with highest probabilities so that the best values are used in the computation process during the speech recognition process, as taught by Gales (Page 1, Col. 2, Ln. 23-34). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Nakagawa et al. with the steps of Zhao and Gales et al. in order to reduce the amount of computation needed to recognize speech.

As per claim 7, Nakagawa et al. as modified by Zhao and Gales et al. teach all the limitations of the method of claim 6. Nakagawa et al. fail to teach the method of claim 6 wherein the recognizing includes: recognizing the input speech vector using the one or more Gaussian distributions with the highest probabilities. Recognizing the input speech vector using the one or more Gaussian distributions with the highest probabilities is well known in the art as taught by Gales et al. (Page 1, Col. 1, Ln. 24-26 and Col. 2, Ln. 19-34). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Nakagawa et al. wherein the recognizing includes recognizing

the input speech vector using the one or more Gaussian distributions with the highest probabilities in order to reduce the amount of computation needed to recognize speech.

As per claim 8, Nakagawa et al. as modified by Zhao and Gales et al. teach all the limitations of the method of claim 1. Nakagawa et al. further disclose the method of claim 1 further comprising: generating the table (Abstract, Ln. 17-18).

As per claim 9, Nakagawa et al. as modified by Zhao and Gales et al. teach all the limitations of the method of claim 8. Nakagawa et al. further disclose the method of claim 8 wherein the generating includes:

identifying a set of means and variances (Abstract, Ln. 9-12); and

identifying one-dimensional Gaussian distributions to be used for recognition using the identified set of means and variances (Abstract, Ln. 9-12).

Nakagawa et al. fail to teach:

determining distance values for the one-dimensional Gaussian distributions, and storing the distance values in the table.

Determining distance values for the one-dimensional Gaussian distributions, and storing the distance values in the table is well known in the art as taught by Gales et al. (Col. 4, Ln. 34-37 and Ln. 56-58: the array is a table). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Nakagawa et al. with the step of Gales et al. in order to effectively use Gaussian distributions to model speech.

As per claim 10, Nakagawa et al. as modified by Zhao and Gales et al. each all the limitations of the method of the method of claim 9. Nakagawa et al. fail to teach the method of claim 9 wherein the identifying a set of means and variances includes: estimating the set of means and variances, and determining a representative set of the means and variances from the estimated set of means and variances. Zhao teaches estimating the set of means and variances (Col. 4, Ln. 1-4), and determining a representative set of the means and variances from the estimated set of means and variances (Col. 4, Ln. 27-31). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Nakagawa et al. with the step of Zhao in order to efficiently use more accurate values in the speech recognition process.

As per claim 11, Nakagawa et al. as modified by Zhao and Gales et al. teach all the limitations of the method of claim 9. Nakagawa et al. fail to teach the method of claim 9 wherein identifying one-dimensional Gaussian distributions includes: determining the one-dimensional Gaussian distributions from different combinations of the means and the variances in the identified set of means and variances. The aforementioned feature is well known in the art as taught by Zhao. Zhao teaches wherein identifying one-dimensional Gaussian distributions includes: determining the one-dimensional Gaussian distributions from different combinations of the means and the variances in the identified set of means and variances (Col. 4, Ln. 1-4 and Ln. 27-30). Therefore, it would have been obvious to one of ordinary skill in the art at the time the

invention was made to modify the method for recognizing speech of Nakagawa et al. with the step of Zhao in order to efficiently use more accurate values in the speech recognition process.

As per claim 12, Nakagawa et al. as modified by Zhao and Gales et al. disclose all the limitations of the method of claim 8. Nakagawa et al. further disclose the method of claim 8 wherein the generating includes:

identifying a set of means and variances (Abstract, Ln. 9-12); and

identifying one-dimensional Gaussian distributions to be used for recognition using the identified set of means and variances (Abstract, Ln. 20-23).

Nakagawa et al. fail to teach:

determining distance values for the one-dimensional Gaussian distributions for each of a plurality of dimensions of a plurality of speech vectors; and

storing the distance values in the table.

The aforementioned feature is well known in the art as taught by Zhao. Zhao teaches determining distance values for the one-dimensional Gaussian distributions for each of a plurality of dimensions of a plurality of speech vectors (Col. 4, Ln. 34-35), and storing the distance values in the table (Col. 4, Ln. 56-58). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Nakagawa et al. with the step of Zhao in order to effectively use Gaussian distributions to model speech.

As per claim 13, Nakagawa et al. as modified by Zhao and Gales et al. teach all the limitations of the method of claim 1. Nakagawa et al. further teach the method of claim 1 wherein the determining an address includes determining a separate address for each of a plurality of dimensions of the input speech vector (Abstract, Ln. 19-26).

As per claim 14, Nakagawa et al. as modified by Zhao and Gales et al. teach all the limitations of the method of claim 13. Nakagawa et al. fail to disclose the method of claim 13 wherein the determining a separate address includes: concatenating a code representing a dimension number and a code representing the Gaussian distribution to form each of the separate addresses. Concatenating a code representing a dimension number and a code representing the Gaussian distribution to form each of the separate addresses is obvious in the art. Since an input value for the input speech vector is selected from a subset of possible quantized input values and a value is determined for the Gaussian distribution, it is obvious to use those already computed values to form a code to access a table. Moreover, it is obvious to concatenate any two values to form a code. Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Nakagawa et al. wherein determining an address includes: concatenating a code representing a dimension number and a code representing the Gaussian distribution to form each of the separate addresses since the obvious way to form a code is to combine values.

3. Claims 2 and 3 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nakagawa et al. (WO 00/46788) in view of Zhao (US Patent 5,193,142) and Gales et al. ("State-Based

Gaussain Selection in Large Vocabulary Continuous Speech Recognition Using HMM's, IEEE Transactions on Speech and Audio Processing, Vol. 7, No. 2, March 1999) as applied to claim 1 above, and further in view of Yamada et al. (US Patent 5,991,442).

As per claim 2, Nakagawa et al. as modified by Zhao and Gales et al. teach all the limitations of the method of claim 1. Nakagawa et al. fail to disclose the method of claim 1 further comprising: quantizing each of a plurality of dimensions of the input speech vector. The aforementioned feature is well known in the art as taught by Yamada et al. (Col. 5, Ln. 51, Ln. 57, Ln. 66-67 and Col. 6, Ln. 19-23). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Nakagawa et al. to further comprise quantizing each of a plurality of dimension of the input speech vector of Yamada et al. in order to create code values.

As per claim 3, Nakagawa et al. as modified by Zhao and Gales et al. teach all the limitations of the method of claim 2. Nakagawa et al. fail to teach the method of claim 2 wherein the quantizing includes: selecting an input value for the input speech vector from a subset of possible quantized input values. The aforementioned feature is well known in the art as taught by Yamada et al (Col. 6, Ln. 19-23). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Nakagawa et al. wherein the quantizing includes selecting an input value for the input speech vector from a subset of possible quantized input values of Yamada et al. in order to create a code value.

As per claims 15 and 16, they are drawn to a system corresponding to the method of claim 1, and they are therefore rejected for similar reasons as set forth in the rejection of claim 1, above.

As per claims 17-19, they are drawn to a system corresponding to the method of claims 2-4 respectively, and are therefore rejected for the same reasons set forth in the rejection of claim 2-4, above.

As per claim 20, it is drawn to a system corresponding to the method of claim 6 and it therefore rejected for the same reason set forth in the rejection of claim 6, above.

As per claims 21-27, they are drawn to a system corresponding to the method of claims 8-14, respectively, and are therefore rejected for the same reasons set forth in the rejection of claims 8-14, above.

As per claim 28, it is a computer-readable medium that corresponds to the method of claims 1 and 2, and is therefore rejected for the same reasons set forth in the rejections of claims 1 and 2, above.

As per claim 29, it is a computer-readable medium that corresponds to the method of claim 1, and is therefore rejected for the same reason set forth in the rejections of claim 1, above.

As per claims 30-33, they are rejected for the same reasons set forth in the rejection of claims 9-12, above.

4. Claims 34, 35, 37 and 44 are rejected under 35 U.S.C. 103(a) as being unpatentable over Yamada et al. (US Patent 5,991,442) in view of Nakagawa et al. (WO 00/46788) and Gales et al. ("State-Based Gaussain Selection in Large Vocabulary Continuous Speech Recognition Using HMM's, IEEE Transactions on Speech and Audio Processing, Vol. 7, No. 2, March 1999).

As per claim 34, Yamada et al. disclose a method for determining multi-dimensional Gaussian distribution likelihood for an input speech vector using a small number of one-dimensional Gaussian distributions, comprising:

receiving an input speech vector having a plurality of dimensions (Col. 7, Ln. 63-64); and

identifying a one-dimensional Gaussian distribution for each dimension of the input speech vector (Col. 8, Ln. 11-15).

Yamada et al. fail to teach:

determining, from a table, probabilities of the one-dimensional Gaussian distributions for the dimensions of the input speech vector using codes representing the one-dimensional Gaussian distributions and numbers representing the dimensions; and

determining the likelihood of a multi-dimensional Gaussian distribution based on the determined probabilities.

The aforementioned features are well known in the art as taught by Nakagawa et al. and Gales et al. Nakagawa et al. teach determining, from a table, probabilities of the one-dimensional Gaussian distributions for the dimensions of the input speech vector using codes representing the one-dimensional Gaussian distributions and numbers representing the dimensions (Abstract, Ln. 9-23). Furthermore, Gales et al. teach determining the likelihood of a multi-dimensional Gaussian distribution based on the determined probabilities (Page 1, Col. 2, Ln. 1-5 and 25-34).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for recognizing speech of Yamada including the steps of determining, from a table, probabilities of the one-dimensional Gaussian distributions for the dimensions of the input speech vector using codes representing the one-dimensional Gaussian distributions and numbers representing the dimensions and determining the likelihood of a multi-dimensional Gaussian distribution based on the determined probabilities for the purpose of reducing the amount of computation in the speech recognition process.

As per claim 35, Yamada et al. as modified by Nakagawa et al. and Gales et al. disclose all the limitations of the method of claim 34. Yamada et al further teach the method of claim 24 wherein the identifying includes: identifying a one-dimensional Gaussian distribution for each dimension of each multi-dimensional Gaussian distribution under consideration (Col. 6, Ln. 19-24 and Col. 8, Ln. 11-15).

As per claim 37, Yamada et al. as modified by Nakagawa et al. and Gales et al. disclose all the limitations of the method of claim 36. Yamada fails to disclose the method of claim 36 wherein the identifying the small number of one dimensional Gaussian distributions includes:

determining the codes representing the one-dimensional Gaussian distributions from codes representing the means and codes representing the variances in the identified set of means and variances.

Determining the codes representing the one-dimensional Gaussian distributions from codes representing the means and codes representing the variances in the identified set of means and variances is obvious in the art. Nakagawa et al. teach storing the means and variances of one-dimensional Gaussian distributions (Abstract, Ln. 9-12). It is obvious to use those already computed values to form a code. Moreover, it is obvious to concatenate any two values to form a code since the obvious way to form a code is to combine values. Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the method for determining multi-dimensional Gaussian distribution likelihood for an input speech vector of Yamada et al. to include the step of determining the codes representing the one-dimensional Gaussian distributions from codes representing the means and codes representing the variances in the identified set of means and variances since the obvious way to form a code is to combine values.

As per claim 44, Yamada et al. teach the computer-readable medium of claim 42. Yamada et al. fail to teach the computer-readable medium of claim 42 wherein the instructions for determining probabilities of the one-dimensional Gaussian distributions include: instructions

for using the one-dimensional Gaussian distribution codes and the dimension numbers as addresses for accessing the table. Instructions for using the one-dimensional Gaussian distribution codes and the dimension numbers as addresses for accessing the table is obvious in the art. Nakagawa et al. teach using quantized values of the input vector as indices into tables carrying information about Gaussian distributions (Abstract, Ln. 14-17). Similarly, a code for the Gaussian distribution may be employed. It is obvious to use available values to form a code to access a table.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the computer-readable medium of Yamada et al. wherein the instructions for determining probabilities of the one-dimensional Gaussian distributions include: instructions for using the one-dimensional Gaussian distribution codes and the dimension numbers as addresses for accessing the table in order to access the probabilities of the Gaussian distributions using available values.

As per claim 36, it is rejected for the same reason set forth in the rejections of claim 9, above.

As per claims 38-41, they are drawn to a system corresponding to the method of claims 34-37 respectively, and are therefore rejected for the same reasons set forth in the rejection of claim 34-37, above.

As per claims 42-43, they are drawn to a computer-readable medium corresponding to the method of claims 38-39 respectively, and are therefore rejected for the same reasons set forth in the rejection of claim 38-39, above.

Conclusion

5. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

US Patent 5,946,656 to Rahim et al. with respect to modeling covariances

US Patent 6,591,235 to Chen et al. with respect to gaussianization

US Patent 5,535,305 to Acero et al. with respect to probability density functions

US Patent 6,539,351 to Chen et al. with respect to high dimensional modeling

US Patent 6,411,930 to Burges with respect to gaussian mixture models

US Patent 6,526,379 to Rigazio et al. with respect to Gaussain distributions and distance

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Kinari Patel whose telephone number is 703-305-8487. The examiner can normally be reached on 9 AM - 5 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 703-305-9645. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

kp

RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER